Polynucleotides and polypeptides coded by said polynucleotides involved in the synthesis of diketopiperazine derivatives

The present invention relates to novel isolated, natural or synthetic polynucleotides and to the polypeptides encoded by said polynucleotides, that are involved in the synthesis of diketopiperazine derivatives, to the vectors comprising said polynucleotides, to the microorganisms transformed with said polynucleotides, to the applications of said polynucleotides and of said polypeptides, and also to processes for synthesizing diketopiperazine derivatives, including cyclodipeptides and diketopiperazine derivatives substituted in the 3- and 6-positions with α,β-unsaturated amino acid side chains.

For the purposes of the present invention, the term "diketopiperazine derivatives" is intended to mean molecules having a diketopiperazine ring (piperazine-2,5-diones or 2,5-dioxopiperazines or 2,5-DKP), substituted in the 3- and the 6-positions with amino acids. In the particular case of cyclic diamino acids (cyclodipeptides or cyclic dipeptides), the substituent groups in the 3- and 6-positions are the amino acid side chains. In the particular case of bisdehydro cyclic diamino acids (bisdehydro cyclic dipeptides), the substituent groups in the 3- and 6-positions are α,β-unsaturated amino acid side chains (figure 1).

Diketopiperazine derivatives constitute a family of compounds essentially produced by microorganisms such as bacteria, yeast, filamentous fungi and lichens. Others have also been isolated from marine organisms, such as sponges and starfish. An example of these derivatives has been demonstrated in humans: cyclo(L-His-L-Pro).

The diketopiperazine derivatives have very varied structures ranging from simple cyclodipeptides to much more complex structures.

The simple cyclodipeptides constitute only a small fraction of the diketopiperazine derivatives, the majority of which have more complex structures in which the main ring and/or the side chains comprise many modifications: introduction of carbon-based, hydroxyl, nitro, epoxy, acetyl or methoxy groups, and also the formation of disulfide bridges or of heterocycles. The formation of a double bond between two carbons is also quite widespread. Certain derivatives, of marine origin, incorporate halogen atoms.

Some examples of amino acids incorporated into the cyclodipeptides are given in table I below:

Table 1

Cyclodipeptide	Organism
Cyclo(Gly-L-Pro)	*Luidia clathrata*
Cyclo(L-Pro-L-Leu)	*Rosellinia necatrix*
Cyclo(L-Ala-L-Val)	*Aspergillus ochraceus*
Cyclo(L-Ala-L-Leu)	*Aspergillus niger*
Cyclo(D-Ala-N-méthyl-L-Leu)	*Beauveria nivea*
Cyclo(L-Pro-L-Val) Cyclo(L-Pro-L-Leu)	*Aspergillus ochraceus* *Rosellinia necatrix*
Cyclo(D-Val-L-Trp)	Aspergillus chevalieri
Cyclo(L-Phe-L-Phe)	*Penicillium nigricans* *Streptomyces noursei*
Cyclo(ΔPhe-ΔLeu) (albonoursin)	*Streptomyces noursei*
Cyclo(L-Pro-L-Tyr)	*Alternaria alternata*
Cyclo(L-Pro-L-Trp)	*Penicillium brevicompactum*
Cyclo(L-Ser-L-Ser)	*Streptomyces orchidaceus*
Cyclo(L-Arg-D-Pro)	*Pseudomonas sp.*
Phenylhistine Roquefortine	*Penicillium roquefortii*
Cyclo(L-Trp-ΔAba)	*Streptomyces spectabilis*
Cyclo(4-methyl-D-Pro-L-Nva)	*Calyx cf. podatypa*
Cyclo(ΔAla-L-Val)	*Pseudomonas aeruginosa*

Very little is known regarding the physiological role of diketopiperazine derivatives. It has been described that the cyclo(ΔAla-L-Val) produced

by *Pseudomonas aeruginosa* could be involved in interbacterial communication signals. Other compounds are described as being involved in the virulence of pathogenic microorganisms or else as binding to iron or as having neurobiological properties.

The diketopiperazine derivatives have proved to be advantageous since the discovery that some of them have biological properties such as, for example, antibacterial, antifungal, antiviral, immunosuppressive or antitumor activities.

Table II below gives some examples of diketopiperazine derivatives having a known biological activity:

Table II

Molecules	Organism	Activity
Ambewelamides A and B	*Usnea sp.*	Cytotoxicity
Aranotin	*Arachniotus aureus*	Antiviral
Bicyclomycin	*Streptomyces sapporonensis*	Antibacterial (inhibition of transcription termination)
Cyclo(Δ-Ala-L-Leu)	*Penicillium sp.* (F70614)	Inhibition of α-glucosidase
Cyclo(N-methyl-Tyr)$_2$	*Streptomyces griseus*	Inhibition of calpain
Cyclo(Trp-Δ-Aba)	*Streptomyces spectabilis*	Inhibition of glutathione-S-transferase
Gliotoxin	*Aspergillus flavus*	Herbicidal, antifungal, antibacterial, antiviral
Haematocin	*Nectria haematococca*	Antifungal
Hyalodendrin	*Penicillium turbatum*	Antibiotic
Mycelianamide	Penicillium sp.	Antibacterial (inhibition of butylcholinesterase)
Phenylhistine	*Aspergillus ustus* (NSC-FO38)	Inhibition of microtubule polymerization
Tan-1496 A, C and E	*Microphaeropsis sp.* (FL-16144)	Inhibition of topoisomerase Antibacterial (Gram+)
Verticillin A	*Gliocladium sp.* (SCF-1168)	Inhibition of induction of the c-fos protooncogene
XR334	*Streptomyces sp.* (X01/4/100)	Inhibition of PAI-I

Although the study of these molecules has become widely developed, little is known regarding their synthesis. It is generally known that, in bacteria and in fungi, these molecules are produced by non-ribosomal biosynthesis. In certain cases, it has been possible to show that the formation of the diketopiperazine ring occurs in molecules which are pre-activated via a thioester bond with an enzyme and

for which the cis-conformation of the peptide bond, which is necessary for the cyclization reaction, is promoted by the presence of proline residues. In other cases, it has been demonstrated that N-alkylation, particularly N-methylation, of the amino acid residues also promotes the cis-conformation of the peptide bond.

Thus, all these studies carried out to date have demonstrated that the primary structure of the precursor molecule, which conditions its conformation, is fundamental for the formation of the diketopiperazine ring to take place and for the process to result in the production of the final diketopiperazine derivative.

However, diketopiperazine derivatives exist which do not contain a proline residue or an N-alkylated residue. By way of example of such derivatives, mention may be made of albonoursin, or cyclo(ΔPhe-ΔLeu), an antibiotic produced by *Streptomyces noursei*. It is known that there exists in *Streptomyces noursei* an enzyme activity which catalyzes the final step of the production of albonoursin, namely the formation of the α,β-unsaturated residues (Gondry et al., Eur. J. Biochem., 2001, 268, 1712-1721). However, this enzyme activity requires a substrate in cyclic form, cyclo[L-Phe-L-Leu), which does not contain a proline residue or any N-alkylated residue, and for which the synthetic pathway is unknown.

Thus, the diketopiperazine derivatives exhibit a very great structural diversity and very varied biological activities which make them advantageous molecules for discovering and developing novel medicinal products.

To do this, it is necessary to be able to have large amounts of these molecules.

Admittedly, pathways for the chemical synthesis of diketopiperazine derivatives have been described, but for the most complex derivatives, the yields are low and the processes can not always be industrialized.

An understanding of the pathways for the

natural synthesis of the diketopiperazine derivatives, particularly that of cyclodipeptides, could enable a reasoned genetic improvement in the producer organisms, and would open up perspectives for substituting or improving the existing processes for synthesis (via chemical or biotechnological pathways) through the optimization of production and purification yields. In addition, modification of the nature and/or of the specificity of the enzymes involved in the biosynthetic pathway for the diketopiperazine derivatives could result in the creation of novel derivatives with original molecular structures and with optimized biological properties.

The present invention falls within this context.

In studying the synthetic pathway for albonoursin, the inventors have demonstrated a polynucleotide (hereinafter referred to as BamH1 polynucleotide (SEQ ID No.5)), comprising four open reading frames, each one encoding a polypeptide responsible for each one of the steps for the synthesis and for the transport of albonoursin, from L-phenylalanine and L-leucine residues, in *Streptomyces noursei* and in heterologous hosts such as *Streptomyces lividans* (see figures 2 and 3).

The inventors have been able to show that:

- the first open reading frame orf1 (albA, SEQ ID No.1) encodes a polypeptide (AlbA, SEQ ID No.6) involved in a cyclodipeptide oxidase (CDO) activity such as that described in Gondry et al., (Eur. J. Biochem., 2001, 268, 1712-1721 (α,β-desaturation);
- the second open reading frame orf2 (albB, SEQ ID No.2) encodes a polypeptide which is translated as two isoforms ($AlbB_1$, SEQ ID No.7 and $AlbB_2$, SEQ ID No.8) required for the activity of the AlbA polypeptide. The two isoforms of AlbB which are expressed in more or less equivalent amounts differ from one another by virtue of the presence of 5 additional amino acids located at the N-terminal end

of $AlbB_1$ and resulting from the use of two different initiation codons. In the case of $AlbB_1$, the initial methionine is eliminated;

- the third open reading frame orf3 (albC, SEQ ID No.3) encodes a polypeptide (AlbC, SEQ ID No.9) which shows no similarity with a peptide synthetase and which is capable of catalyzing the condensation of two amino acid residues so as to form a cyclic dipeptide. For example in *Streptomyces noursei*, AlbC catalyzes the condensation of an L-phenylalanine and of an L-leucine or of two L-phenylalanines, so as to form the cyclic dipeptide cyclo(L-Phe-L-Leu), which is a precursor required for the formation of albonoursin and the cyclic dipeptide cyclo(L-Phe-L-Phe). In this particular case, AlbC catalyzes the cyclization of amino acid residues which are neither a proline nor an N-alkylated residue, and
- the fourth open reading frame orf4 (albD, SEQ ID No.4) encodes a polypeptide (AlbD, SEQ ID No.10) which is not directly involved in the succession of reactions resulting in the formation of α,β-unsaturated diketopiperazine derivatives from amino acids, but is probably involved in the mechanism of transport of said derivatives.

The inventors have thus shown that, for the synthesis of α,β-unsaturated diketopiperazine derivatives, only the three open reading frames albA, albB and albC are absolutely necessary, particularly for the synthesis of albonoursin in *Streptomyces noursei*.

Thus, a subject of the invention is an isolated, natural or synthetic polynucleotide characterized in that it comprises at least the three open reading frames albA, albB and albC corresponding, respectively, to the sequences SEQ ID No.1, SEQ ID No.2 and SEQ ID No.3.

This polynucleotide encodes the enzymes required for the synthesis of α,β-unsaturated diketopiperazine derivatives.

According to an advantageous embodiment of the invention, said polynucleotide also comprises the open reading frame albD corresponding to the sequence SEQ ID No.4.

This polynucleotide encodes the enzymes required for the synthesis of α,β-unsaturated diketopiperazine derivatives and for their transport and for their secretion.

According to a particular form of the invention, the polynucleotide corresponds to the sequence SEQ ID No.5. This polynucleotide (BamH1 polynucleotide) contains the four open reading frames albA, albB, albC and albD, and therefore encodes the enzymes required for the synthesis of α,β-unsaturated diketopiperazine derivatives and for their transport and for their secretion.

A subject of the invention is also an isolated, natural or synthetic polynucleotide characterized in that it comprises at least one of the three open reading frames albB, albC and albD corresponding, respectively, to the sequences SEQ ID No.2, SEQ ID No.3 and SEQ ID No.4.

Thus, the polynucleotide comprising the open reading frame albC (SEQ ID No.3) encodes an enzyme which allows the cyclization of two amino acids, which may be identical or different, so as to form a cyclic dipeptide. The polynucleotide comprising the open reading frames albC and albD (SEQ ID No.3 and SEQ ID No.4) encodes the enzymes which allow, firstly, the cyclization of two amino acids, which may be identical or different, so as to form a cyclic dipeptide and, secondly, the transport of said dipeptide.

The subject of the invention is also an isolated, natural or synthetic polynucleotide corresponding to any one of the sequences SEQ ID No.2 (albB, 318 nucleotides), SEQ ID No.3 (albC, 720 nucleotides) or SEQ ID No.4 (albD, 834 nucleotides).

A subject of the invention is also fragments of

the polynucleotides as defined above. The term "fragment" is intended to mean any sequence of at least 15 nucleic acids.

The polynucleotide according to the invention can be obtained from DNA libraries, particularly from microorganism DNA libraries, very particularly from a *Streptomyces noursei* DNA library. The polynucleotide of the invention can also be obtained by means of a polymerase chain reaction (PCR) carried out on the total DNA of *Streptomyces noursei*. The polynucleotides according to the invention can be obtained by RT-PCR carried out on the total RNA of *Streptomyces noursei*.

A subject of the invention is also a vector into which is inserted any one of the polynucleotides described above. Thus, the vector of the invention can comprise the polynucleotide comprising the three or the four open reading frames albA, albB, albC and/or albD, corresponding, respectively, to the sequences SEQ ID No.1, SEQ ID No.2, SEQ ID No.3 and/or SEQ ID No.4, the polynucleotide comprising at least one of the three open reading frames albB, albC or albD corresponding, respectively, to the sequences SEQ ID No.2, SEQ ID No.3 and SEQ ID No.4, any one of the polynucleotides corresponding to at least one of the three open reading frames albB, albC or albD corresponding, respectively, to the sequences SEQ ID No.2, SEQ ID No.3 and SEQ ID No.4, the BamH1 polynucleotide corresponding to the sequence SEQ ID No.5 or else a fragment of said polynucleotides.

The vector used may be any known vector of the prior art. As vectors that can be used according to the invention, mention may in particular be made of plasmids, cosmids, bacterial artificial chromosomes (BACs), integrative elements of actinobacteria, viruses or else bacteriophages.

Said vector may also comprise any regulatory sequences required for the replication of the vector and/or the expression of the polypeptide encoded by the polynucleotide (promoter, termination sites, etc.).

A subject of the invention is also the use of at least one of the polypeptides as defined above, or of one of its fragments, as a probe for detecting corresponding sequences in other organisms or as a primer for amplifying such sequences.

When they are primers, said polynucleotides or said fragments also include antisense sequences.

One of the preferred uses of the probes or primers described above is the investigation of polynucleotide sequences homologous to the sequences of the open reading frames albA, albB, albC or albD in other organisms, in order in particular to demonstrate novel synthetic pathways for diketopiperazine derivatives.

A subject of the invention is also an isolated, natural or synthetic polypeptide characterized in that it comprises at least any one of the sequences SEQ ID No.7 to SEQ ID No.10, corresponding, respectively, to the polypeptides $AlbB_1$, $AlbB_2$, AlbC or AlbD.

A subject of the invention is also an isolated, natural or synthetic polypeptide characterized in that it corresponds to any one of the sequences SEQ ID No.7 ($AlbB_1$), SEQ ID No.8 ($AlbB_2$), SEQ ID No.9 (AlbC) or SEQ ID No.10 (AlbD).

The invention also relates to the polypeptides encoded by any one of the polynucleotides of the invention, particularly any one of the polynucleotides chosen from any one of the sequences SEQ ID No.2 (albB), SEQ ID No.3 (albC) or SEQ ID No.4 (albD).

Advantageously, the polypeptides according to the invention can be either isolated from microorganisms (*Streptomyces noursei*) for example, or obtained by chemical synthesis or else by biotechnological means, from the polynucleotides of the invention, such as, for example, from modified microorganisms which do not normally express said polypeptides.

A subject of the invention is also an isolated

polypeptide, the sequence of which is substantially homologous to at least one of the sequences SEQ ID No.7 to SEQ ID No.10, as defined above.

It is considered here that a polypeptide has a substantially homologous sequence when its amino acid sequence exhibits at least 80% similarity with the amino acid sequence of at least one of the sequences SEQ ID No.7 to SEQ ID No.10 and when the polypeptide has conserved its initial activity.

The expression "80% similarity between a polypeptide P and the sequences SEQ ID No.7 to 10" is intended to mean that, when the two polypeptides are aligned, 80% of the amino acids of P are identical to the corresponding amino acid of the sequences SEQ ID No.7 to 10 or are replaced with an amino acid of the same group.

The expression "amino acid of the same group" is intended to mean an amino acid having substantially identical chemical properties. In particular, this term is intended to mean amino acids having substantially the same charge and/or the same size and/or the same hydrophilicity or hydrophobicity and/or the same aromaticity.

Such amino acid groups include in particular:

(i) glycine, alanine
(ii) isoleucine, leucine, valine
(iii) tryptophan, tyrosine, phenylalanine
(iv) aspartic acid, glutamic acid
(v) arginine, lysine, histidine
(vi) serine, threonine.

Other substitutions can be envisioned, in which an amino acid is replaced with another amino acid that is comparable but not natural (hydroxyproline, norleucine, ornithine, citrulline, cyclohexylalanine, dextrorotatory amino acids, etc.).

A subject of the invention is also the use of the polynucleotides or the vectors of the invention as described above, for synthesizing polypeptides

corresponding to the sequences SEQ ID Nos.7 to 10.

A subject of the invention is also the use, particularly *in vitro*, of the polypeptides according to the invention, alone or in combination, for preparing cyclodipeptides and/or diketopiperazine derivatives substituted in the 3- and 6-positions with α,β-unsaturated amino acid side chains, particularly albonoursin.

A subject of the invention is also the use of the polypeptides of the invention, alone or in combination, for modifying the pharmacological activity of a biological molecule by modifying its structure, for example by dehydrogenation of side chains, particularly of amino acid side chains, or by cyclization, particularly of peptide molecules.

A subject of the invention is also a modified biological system into which at least one polynucleotide according to the invention or at least one vector according to the invention has been introduced.

Such a biological system may be any known heterologous expression system using prokaryotes or eukaryotes as hosts. By way of example, mention may be made of a microorganism, for instance a bacterium such as *Escherichia coli* or *Streptomyces lividans*, or animal or insect cells.

A subject of the invention is also a modified *in vitro* acellular system into which at least one polynucleotide according to the invention or at least one vector according to the invention has been introduced.

A subject of the invention is also the use of at least one polynucleotide according to the invention and/or of at least one vector according to the invention, for preparing a modified biological system, it being possible for said system to be a microorganism, for instance a bacterium such as *Escherichia coli* or *Streptomyces lividans*, or any known heterologous expression system using prokaryotes or

eukaryotes as hosts, or else a modified *in vitro* acellular system.

The introduction of the polynucleotide and/or of the vector according to the invention into the host modified biological system can be carried out by any known method, such as, for example, transfection, infection, fusion, electroporation, microinjection or else biolistics.

A subject of the invention is also the use of at least one modified biological system or of at least one modified *in vitro* acellular system, as defined above, for preparing cyclodipeptides and/or diketopiperazine derivatives substituted in the 3- and 6-positions with α,β-unsaturated amino acid side chains, particularly albonoursin.

The biological systems are suitable for the synthesis, with a good yield, of the cyclodipeptides and of the α,β-unsaturated diketopiperazine derivatives as defined above.

When it is a microorganism, the modified biological system may also optionally allow the secretion of the diketopiperazine derivative according to the invention into a culture medium, making it easier to extract it and purify it. The presence of AlbD in the biological systems such as microorganisms constitutes an advantageous step for the industrial process by facilitating the extraction and purification of the derivative, which is thus secreted into the culture medium.

A subject of the invention is also a method for the *in vitro* synthesis of a cyclodipeptide, characterized in that:

(1) two amino acids, which may be identical or different, are brought into contact, under suitable conditions, with AlbC (SEQ ID No.9) and

(2) the cyclodipeptide obtained is purified.

A subject of the invention is also a method for the *in vitro* synthesis of a diketopiperazine derivative substituted in the 3- and 6-positions with

α,β-unsaturated amino acid side chains, characterized in that:

(1) two amino acids, which may be identical or different, are brought into contact, under suitable conditions, with AlbC (SEQ ID No.9) and

(2) the cyclodipeptide obtained in step (1) is brought into contact with AlbA (SEQ ID No.6), AlbB1 (SEQ ID No.7) and AlbB2 (SEQ ID No.8), and then the α,β-unsaturated diketopiperazine derivative obtained is purified. The method may also include, in step (2), the polypeptide AlbD (SEQ ID No.10). The method may optionally comprise, between step (1) and step (2), an additional step for purification of the cyclodipeptide obtained in step (1).

This method can, of course, be carried out in a single step in which two amino acids, which may be identical or different, are brought into contact, under suitable conditions, with AlbA (SEQ ID No.6), AlbB1 (SEQ ID No.7), AlbB2 (SEQ ID No.8) and AlbC (SEQ ID No.9), optionally AlbD (SEQ ID No.10), and the α,β-unsaturated diketopiperazine derivative obtained is purified.

The term "suitable conditions" is preferably intended to mean the conditions under which the following are incubated:

- the polypeptides (AlbA, AlbB, AlbC and/or AlbD) at concentrations of between 0.1 nM and 10 μM, preferably of between 10 nM and 1 μM;
- in the presence of amino acids, which may be identical or different, at a concentration of between 0.1 mM and 100 mM, preferably of between 1 mM and 10 mM;
- in a 0.1 M Tris-HCl buffer, having a pH of between 6.8 and 8.0, at a temperature of between 28°C and 40°C, for a period of time of between 2 hours and 48 hours.

A subject of the invention is also a method for synthesizing a cyclodipeptide, characterized in that:

(1) a biological system comprising at least the polynucleotide SEQ ID No.3 (albC, encoding AlbC) is brought into contact, under conditions suitable for culturing said chosen biological system, and

(2) the cyclodipeptide obtained is purified. The biological system may also comprise the polynucleotide SEQ ID No.4 (encoding AlbD).

A subject of the invention is also a method for synthesizing a diketopiperazine derivative substituted in the 3- and 6-positions with α,β-unsaturated amino acid side chains, characterized in that:

(1) a biological system comprising the polynucleotide corresponding to the sequences SEQ ID Nos.1 to 3 (encoding AlbA, AlbB, and AlbC) is brought into contact, under conditions suitable for culturing said chosen biological system, and

(2) the α,β-unsaturated diketopiperazine derivative obtained is purified. The biological system may also comprise the polynucleotide corresponding to the sequence SEQ ID No.4 (encoding AlbD).

The expression "conditions suitable for culturing said chosen biological system" is understood to mean that the method is carried out under the conditions for culturing the chosen biological system, including a suitable culture medium containing a large excess of amino acids. For example, if the biological system is a microorganism, such as, for example, *Escherichia coli*, the suitable conditions are those commonly used for culturing this bacterium. The same is true for *Streptomyces lividans* or if the biological system is a eukaryotic cell.

According to the methods of the invention, the amino acids, which may be identical or different, are present in an amount of between 0.1 mM and 100 mM, preferably of between 1 mM and 10 mM.

Similarly, according to the methods of the invention, the polypeptides AlbA, AlbB, AlbC and AlbD are present in an amount of between 0.1 nM and 10 μM, preferably of between 10 nM and 1 μM.

The purification of the cyclodipeptides and of the α,β-unsaturated diketopiperazine derivatives can be carried out directly from syntheses *in vivo* or *in vitro* by means of liquid-phase extraction techniques or by means of precipitation, or thin-layer or liquid-phase chromatography techniques, in particular reverse-phase HPLC, or any method suitable for purifying peptides, one known to those skilled in the art.

The methods of the invention may be carried out in any suitable biological system, particularly in a host such as, for example, a microorganism, for instance a bacterium such as *Escherichia coli* or *Streptomyces lividans*, or any known heterologous expression system using prokaryotes or eukaryotes as hosts, or even *in vitro* acellular systems.

Besides the above provisions, the invention also comprises other provisions which would emerge from the following description, which refers to examples of implementation of the invention and also to the attached drawings, in which:

- Figure 1 represents the chemical structures of a diketopiperazine ring (A) and of albonoursin (B).

- Figure 2 represents a diagram of the genomic region of *Streptomyces noursei* including the gene cluster required for the synthesis of albonoursin, in which orf1 corresponds to albA, orf2 corresponds to albB, orf3 corresponds to albC, and orf4 corresponds to albD.

- Figure 3 represents the presumed synthetic pathway for albonoursin *Streptomyces noursei*.

- Figure 4 represents certain plasmid constructs prepared in order to introduce the various open reading frames (or orfs) into *Escherichia coli* or *Streptomyces lividans*.

- Figure 5 represents the results of the analyses of the culture media for *Streptomyces lividans*, transformed by introduction of the plasmids pSL128(A), pUWL201 (B) and pSL129 (C).

- Figure 6 represents the results of the

analyses of the culture media for *Streptomyces lividans*, transformed by introduction of the plasmids pSL168 and pSL159, or for nontransformed *Streptomyces lividans*:

A: *S. lividans*[pSL168], incubated in the presence of CDO;

B: *S. lividans*[pSL168], incubated in the absence of CDO;

C: *S. lividans*[pSL159] in the absence or in the presence of CDO;

D: *Streptomyces lividans* TK21 incubated in the presence of CDO.

- Figure 7 represents the sodium dodecyl sulfate/polyacrylamide gel (SDS-PAGE) analysis of the expression of AlbA and $AlbB_1$ and $AlbB_2$ from the polynucleotide albA-albB in *E. coli* BL21(DE3)-plysS. Staining with Coomassie brilliant blue R-250. Lane S: molecular weight standard, lane 1: total cytoplasmic extract (approximately 10 µg), lane 2: enzyme fraction purified on an Ni-sepharose column (approximately 10 µg).

The following examples illustrate the invention but in no way limit it.

EXAMPLE 1: Isolation of the polynucleotide of the invention in *Streptomyces noursei*

A polynucleotide containing all the genetic information required for the biosynthesis of albonoursin in *Streptomyces noursei* was isolated from the total genome of this microorganism by an approach based on gene amplification by PCR.

- Obtaining partial peptide sequences of the cyclodipeptide oxidase:

Partial peptide sequence information on the enzyme which catalyzes, in *Streptomyces noursei*, the conversion of the cyclodipeptide cyclo(L-Phe-L-Leu) to albonoursin, called cyclodipeptide oxidase (CDO), was obtained by direct sequencing, by the Edman method, of polypeptides derived from trypsin hydrolysis of the purified enzyme according to the protocol described in

M. Gondry et al. (Gondry et al., 2001, mentioned above). After separation of the enzyme fraction constituents by 15% polyacrylamide gel electrophoresis and staining of the proteins with Coomassie blue, a gel band containing the protein of molecular mass of approximately 21 000 daltons is cut out and incubated in 1 ml of 50 mM Tris-HCl buffer, pH 8, in the presence of trypsin (relative trypsin/substrate concentrations = 1/50), for 20 hours at 37°C. The polypeptides obtained are then separated by reverse-phase high performance liquid chromatography (HPLC) (μRPC C_2/C_{18} SC21/10 column, Pharmacia) with a linear gradient of 0% to 76% of acetonitrile in 62 minutes (solvent: 0.1% trifluoro-acetic acid; flow rate: 1 ml/min). Each of the separated polypeptides is then purified by gel permeation chromatography on a superdex peptide PC32/30 column (Pharmacia) equilibrated in buffer containing 30% of acetonitrile and 0.1% of trifluoroacetic acid. Three of the polypeptides obtained were finally analyzed by automatic sequencing by the Edman method (model 477A sequencer, Applied Biosystems) and by MALDI-TOF mass spectrometry. Table 1 gives the peptide sequences obtained and also the nucleotide sequences which were deduced therefrom.

Table I:

Peptide sequence	Deduced nucleotide sequence
EPVDDALIEQLLEAMLAAPT (SEQ ID N°11)	GARCCSGTSGACGACGC (oligo1f) (SEQ ID N°14)
	GCGTCGTCSACSGGYTC (oligo1r) (SEQ ID N°15)
NEVVNYEXWGNR (SEQ ID N°12)	AACGARGTSGTSAACTACGA (oligo2f) (SEQ ID N°16)
	TCGTAGTTSACSACYTCGTT (oligo2r) (SEQ ID N°17)
QAXSFMVVR (SEQ ID N°13)	CAGGCSTGGWSSTTCATGGT (oligo3f) (SEQ ID N°18)
	ACCATGAASSWCCASGCCTG (oligo3r) (SEQ ID N°19)

X: undetermined amino acid (R = A or G; S = C or G; Y = C or T and W = A or T).

Comparison of the experimental and theoretical

masses of the polypeptide corresponding to the sequence SEQ ID No.13 makes it possible to identify a tryptophan residue in the 3-position. The sequences underlined and in bold were used to design 6 sense (1f, 2f and 3f) and antisense (1r, 2r and 3r) degenerate oligonucleotides, according to the typical codon use in *Streptomyces*. In the absence of the information regarding the respective position of the polypeptides in the protein sequence, a combination of the six oligonucleotides was used for the cloning. Since the PCR conditions tested resulted in too great a number of amplified nucleotide fragments, a reverse transcription (RT-PCR) strategy was developed.

- Amplification of an oligonucleotide fragment by RT-PCR:

The total RNA of *Streptomyces noursei* was extracted from a 24-hours culture in medium 5 (ATCC medium) according to the protocol described by Kieser et al. (Practical *Streptomyces* Genetics. The John Innes Foundation Norwich, U.K. (2000)). An additional treatment with DNase I makes it possible to completely eliminate the DNA. The degenerate oligonucleotides (sense and antisense) were synthesized by Sigma Genosys Ltd and the RT-PCR was carried out using the Titan™ One Tube RT-PCR kit (Boehringer Mannheim) according to the standard instructions, with 1 µg of total RNA for each reaction. The reverse transcription is carried out at 50°C for 30 minutes, and the PCR conditions are then as follows: initial denaturation at 97°C for 4 min, followed by 45 cycles of 1 min at 95°C, 1 min at 50°C and 1 min at 68°C, and the final polymerization reaction at 68°C for 10 minutes. The reaction products are analyzed by 1.5% agarose gel electrophoresis, and then purified (DNA and Gel Band purification kit, Pharmacia).

The RT-PCR with the six oligonucleotide combinations resulted in the amplification of a single fragment of approximately 400 base pairs, with the oligonucleotides 3f and 2r. This fragment was cloned

into the vector pGEM-T easy vector and sequenced, making it possible to confirm that the oligonucleotide primers used are indeed in the same reading frame. This nucleotide fragment was used as a probe for screening a *Streptomyces noursei* genomic DNA library, prepared in the cosmid pWED1.

- Construction of a *Streptomyces noursei* genomic DNA library:

The genomic DNA (2.5 µg), extracted from *Streptomyces noursei* according to standard procedures (Kieser et al., mentioned above; J. Sambrook et al., Molecular cloning: a laboratory manual. Cold Spring Harbor Laboratory Press. Cold Spring Harbor, N.Y. (1989)) was partially digested with 0.33 U of BamHI, resulting in DNA fragments of approximately 35 to 45 kb. These fragments are introduced by ligation into the cosmid vector pWED1 digested with BamHI and dephosphorylated beforehand. The ligation product was encapsulated *in vitro* in lambda phages (Packagene Lambda DNA packaging system, Promega) and introduced by transfection into the *E. coli* (SURE) strain.

- Screening of the *Streptomyces noursei* genomic DNA library:

The nucleotide fragment amplified by RT-PCR was then labeled by random priming with [α-^{32}P]-dCTP using the T7 Quick Prime kit (Pharmacia), and used as a probe for screening the library. Approximately 2000 clones were tested by colony hybridization according to the standard method (J. Sambrook et al., mentioned above) and 12 clones were selected. The corresponding cosmids (referred to as pSL110 to pSL121) were extracted, digested with BamHI and analyzed by the Southern blotting technique using the RT-PCR fragments as a probe. This probe made it possible to isolate a 3.8 kb nucleotide fragment common to all the cosmids and also present in the *Streptomyces noursei* genomic DNA digested with BamHI. This BamHI fragment was isolated from the cosmid pSL117 and cloned into the vector pBC SK$^+$, to give the vector pSL122 (definition of the

vectors used: cf. table II).

EXAMPLE 2: Analysis of the sequence of the polynucleotide of the invention

The automatic sequencing of the polynucleotides of the invention was carried out on an ABI PRISM Genetic Analyzer (Perkin Elmer) using the DYEnamic ET terminator cycle kit (Pharmacia) or by the company Genome Express. Computer analysis of the sequences and comparisons with the databanks were carried out with the Frame (Bibb, M.J. et al., Gene, 30, 157-166 (1984)), BLAST and FASTA (Altschul, S.F. et al., Nucleic Acids Res., 25, 3389-3402 (1997); Pearson, W.R., Methods in Enzymology, 183, 63-98 (1990)) programs.

Analysis of the BamHI polynucleotide (SEQ ID No.5) by means of the FRAME program reveals four complete open reading frames, referred to as orf1 to orf4 (albA to albD, SEQ ID Nos.1 to 4) transcribed in the same direction, and an open reading frame (1119 bp) the end of which is truncated, referred to as orf5 (see figure 2). The orf5 translation product exhibits a very high degree of similarity with the N-terminal portion of an NADP-specific glutamate dehydrogenase from *Streptomyces coelicolor* (78% identity and 86% similarity according to the BLAST program).

The first of the open reading frames, orf1 (albA, SEQ ID No.1), contains the nucleotide sequence of the fragment amplified by RT-PCR, and the deduced peptide sequence indeed contains the sequence of the 3 trypsin peptides initially isolated. The product of orf1 corresponds, consequently, to the enzymatic protein of approximately 21 kDa in mass, isolated and purified from *Streptomyces noursei* (M. Gondry et al., mentioned above). Consequently, this gene is clearly involved in the biosynthesis of albonoursin and will be referred to as albA.

Analysis of the sequence of albA (orf1) indicates that 3 codons, two GUGs and one AUG, could be

considered as initiation codons for the translation of albA, which would result in proteins of 219, 204 or 196 amino acids. Since attempts to determine the N-terminal peptide sequence failed, due to the presence of a post-translational modification on this end, the longest sequence (657 nucleotides) was selected for albA (SEQ ID No.6) (the first initiation codon is located at a distance of 20 nucleotides from the end of the BamHI fragment which, consequently, does not contain the promoter region which must be located further upstream of this gene).

Comparison of the peptide sequence deduced from albA (AlbA, SEQ ID No.6) with the databases showed a maximum degree of similarity with an NADH oxidase from *Archaeoglobus fulgidus* (32% identity and 52% similarity according to the BLAST program) and the search for conserved domains indicates that it has a large nitroreductase-type domain (pfam00881, 151 amino acids).

Orf2 (albB, SEQ ID No.2), which is contiguous with albA, but with a reading frame shift, also exhibits typical Streptomyces codon use. albB is translated as two isoforms ($AlbB_1$, SEQ ID No.7 and $AlbB_2$, SEQ ID No.8) that are required for the activity of the AlbA polypeptide, according to the initiation codon AUG or GUG, taken into consideration for orf2. The two isoforms of AlbB, which are expressed in reasonably equivalent amount, differ by virtue of the presence of 5 additional amino acids located at the N-terminal end of $AlbB_1$ and that result from the use of two different initiation codons. In the case of $AlbB_1$, the initial methionine is eliminated.

The two possibilities are compatible with the analysis of the sequence using the FRAME program. The BLAST and FASTA programs reveal no particular homology between the peptide sequence deduced from orf2 and the proteins of the databanks.

Similarly, searches in the databanks carried out based on the polypeptide sequences deduced,

respectively, from orf3 (albC, SEQ ID No.3) and orf4 (albD, SEQ ID No.4) reveal no significant homology with a protein of known function. orf3 begins with an ATG initiation codon and encodes a polypeptide of 239 amino acids, AlbC (SEQ ID No.9), which exhibits a low degree of similarity with two hypothetical proteins of unknown function: Rv2275 from *Mycobacterium tuberculosis* (34% identity and 53% similarity according to the BLAST program) and YvmC from *Bacillus subtilis* (29% identity and 46% similarity according to BLAST). Orf4 encodes a 277 amino acid protein, AlbD (SEQ ID No.10), which comprises a transmembrane domain, as indicated by the analysis of its sequence with the TMHMM program (Krogh, A. et al., J. Mol. Biol. 305, (2001)), and exhibits a weak homology with a transmembrane protein of unknown function from *Streptomyces coelicolor* (54% identity and 67% similarity according to BLAST).

EXAMPLE 3: Cloning of the polynucleotides albA, albB, albC and albD and construction of the expression vectors

The methods for extraction and preparation of DNA, for transformation of the *Escherichia coli* and *Streptomyces lividans* TK 21 strains, and for preparation of the protoplasts were carried out according to the standard protocols described by Sambrook et al. (mentioned above) and Kieser et al. (mentioned above).

All the plasmid and cosmid vectors prepared for manipulating the polynucleotides which are the subject of the present invention are given in table II (see also figure 4).

Table II: Strains and vectors used

Bacteria	Properties	Source/reference
E. coli DH5α	standard strain for cloning	Invitrogen
E. coli SURE	strain used for cosmid libraries	Stratagene
Streptomyces lividans TK21	*Streptomyces* strain for cloning the alb genes	Hopwood, D. A. et al. J. Gen. Microbol. 129, 2257-2269 (1983).

Streptomyces noursei ATCC 1145	wild-type strain producing albonoursin	ATCC
Vectors		
pGEM-T easy	Vector for cloning the PCR products, Amp^R	Promega
pWED1	Cosmid vector derived from pWE15, Amp^R	Gourmelen, A. et al., Antimicrobial Agents Chemotherapy, 42, 2612-2619 (1998)
pBC SK^+	Cloning vector, Cm^R	Stratagene
pHP45 Ωaac	Plasmid used at source for the cassette Ωaac, Amp^R, Apr^R	Blondelet-Rouault, M. H. et al., Gene, 190, 315-317 (1997).
pUWL201	*E. coli/Streptomyces* shuttle vector, contains the ErmeE* promoter for the expression of cloned genes in *Streptomyces*. Amp^R, $Thio^R$	Doumith, M., et al., Mol. Gen. Genet. 264, 477-485 (2000)
pET-28a	Expression vector	Novagen
pSL117	Cosmids from the *Streptomyces noursei* library, used for the RT-PCR products	
pSL122	3.8 kb BamHI fragment cloned into pBC SK+	
pSL127	pSL122 derivative, comprising an internal deletion of the ApaI fragment, eliminating orf2, orf3, orf4 and orf5	
pSL128	3.8 kb BamHI fragment of pSL122, cloned into pUWL201, under the control of Erme*p	
pSL129	3.8 kb BamHI fragment of pSL122, cloned into pUWL201, containing the insertion in the opposite direction to pSL128	

pSL138	pSL122 derivative, comprising an internal deletion of the EcoRI fragment, eliminating orf3, orf4 and orf5, and with insertion of the Ωaac cassette	
pSL140	Psl122 derivative, comprising an internal deletion of the NdeI/EcoRV fragment eliminating orf5, and with the insertion of the Ωaac cassette	
pSL142	Asp718/Klenow/BamHI fragment of pSL138 (containing orf1, orf2 and the Ωaac cassette) cloned into pUWL201	
pSL144	Asp718/Klenow/BamHI fragment of pSL140 (containing orf1 to orf4 and the Ωaac cassette) cloned into pUWL201	
pSL145	pSL122 derivative with internal deletion of the EcoRI fragment, eliminating orf3, orf4 and orf5	
pSL150	PCR product from amplification of orf1-orf2, cloned into the expression vector pET-28a	
pSL157	PCR product from amplification of orf4, cloned into pGEM-T easy	
pSL159	Pst1/Klenow/BamHI fragment of pSL157 cloned into pUWL201	
pSL165	PCR product from amplification of orf3, cloned into pGEM-T easy	
pSL166	PCR product from amplification of orf3 + orf4, cloned into pGEM-T easy	
pSL167	PstI/Klenow/BamHI fragment of pSL166 cloned into pUWL201	
pSL168	PstI/Klenow/BamHI fragment of pSL165 cloned into pUWL201	

pSL117 is a cosmid containing the *Streptomyces noursei* genomic DNA library.

pSL122 contains the 3.8 kb BamHI polynucleotide (SEQ ID No.5) from *Streptomyces noursei* which is the subject of the invention, cloned into the cloning vector pBC SK^+. pSL127 and pSL145 were constructed, respectively, by digestion of pSL122 with ApaI or

EcoRI, and religation.

The BamHI polynucleotide was also cloned into the *E. coli/Streptomyces* shuttle vector pUWL201, in the orientation appropriate for having all the genes under the control of the ermE* promoter (pSL128) or else in the opposite orientation (pSL129).

pSL142 and pSL144 were constructed in 2 steps: pSL122 was first digested with EcoRI and the Klenow enzyme, or NdeI and the Klenow enzyme, and then a ligation between these fragments and the Ωaac cassette digested with HindIII-Klenow resulted in the plasmids pSL138 and pSL140. These plasmids were then digested with Asp718, Klenow and BamHI, and the fragments obtained containing orf1 (albA, SEQ ID No.1), orf2 (albB, SEQ ID No.2) and the Ωaac cassette, for the first, and orf1 to orf4 (albA to albD) and the Ωaac cassette, for the second, are cloned into the vector pUWL201 digested with XbaI-Klenow-BamHI.

orf3 (albC, SEQ ID No.3), orf4 (albD, SEQ ID No.4) and (orf3+orf4) were amplified by PCR using the following primers:

for orf3

sylv24 : (SEQ ID N°20) :

5'-CGGCTGCAGGAGAAGGGAGCGGACATATGCTTGCAGGCTTAGTTCCC-3', (PstI site underlined);

sylv22 : (SEQ ID N°21) :

5'-CGGTCCCGTGGATCCAAGCTTCTAGGCCGCGTCGGCCAGCTC-3', (BamHI site underlined);

for orf4

sylv19 : (SEQ ID N°22) :

5'-GAGCGGGATCCTGCAGTGTCATGGGGAGGACAGGAC-3', (PstI site underlined);

sylv18 : (SEQ ID N°23) :

5'-CGATCACGTGGATCCAAGCTTGCCAATCCTGTACGCGATTT-3', (BamHI site underlined);

for (orf3+orf4):sylv24 and sylv18.

orf2 and orf3 are separated only by 37 nucleotides, a synthetic ribosome-binding site was included in sylv24 in order to ensure correct

translation of orf3. The fragments amplified by PCR were then cloned into the vector pGEM-T easy (Promega), to give pSL165, pSL157 and pSL166. The PstI-BamHI fragments obtained from these three plasmids were then cloned into the vector pUWL201 digested with PstI-BamHI, to give pSL168, pSL159 and pSL167.

EXAMPLE 4: Production of the diketopiperazine derivative cyclo(ΔPhe-ΔLeu) (albonoursin) in a heterologous host, *Streptomyces lividans*, transformed by introduction of the BamHI polynucleotide (SEQ ID No.5)

Streptomyces lividans TK21 protoplasts were transformed with pSL128 (containing the BamHI polynucleotide (SEQ ID No.5)), pSL129 or pUWL201 (control) according to the standard protocols described by Sambrook et al. (mentioned above) and Kieser et al. (mentioned above). The three strains were cultured in M5 medium (ATCC medium), a rich medium containing a large excess of amino acids, for 3 days at a temperature of between 28 and 30°C. The supernatants from the cultures of the 3 transformed strains were analyzed by reverse-phase HPLC under the following conditions: the culture supernatant (500 µl) was filtered (ultrafree-MC 10 kDa, Millipore) and injected directly onto HPLC (Vydac C18 column (4.6×250 mm): flow rate: 1 ml/min; elution: linear gradient of 0 to 45% of acetonitrile in 0.1% of trifluoroacetic acid in 45 minutes). The elution was monitored using a multi-wavelength detector for between 200 and 600 nm.

The production of albonoursin in 2 stereoisomeric forms (2 peaks at 38.3 min and 40.5 min; λ_{max} = 318 nm; m = 256.4 Da) is detected in *S. lividans*[pSL128] (figure 5A).

In the control strains *S. lividans*[pUWL201] and *S. lividans*[pSL129] (figures 5B and 5C), no production of albonoursin is detected.

The BamHI polynucleotide (SEQ ID No.5) from *Streptomyces noursei* contains all the genetic information for the production of albonoursin.

EXAMPLE 5: Demonstration of the function of the polynucleotides albA and albB by visualization of the conversion of cyclo(L-Trp-L-Trp) to cyclo(ΔTrp-ΔTrp) in a Petri dish using a heterologous host, *Escherichia coli*, transformed by insertion of the polynucleotides albA and albB.

A rapid test in a Petri dish was developed in order to directly detect the conversion of cyclodipeptides to bisdehydro cyclodipeptides on isolated colonies. This test is based on the conversion of the cyclodipeptide, which is colorless in solution, cyclo(L-Trp-L-Trp), to a yellow and insoluble product, cyclo(ΔTrp-ΔTrp) (λ_{max} = 367 nm and 450 nm), which gives the colonies exhibiting the CDO activity a bright yellow color.

The *E. coli* strains transformed with the plasmids prepared according to the protocol explained in detail in example 3 were tested directly on a dish of LB medium containing 0.5 mM of cyclo(L-Trp-L-Trp).

After incubation for 16 hours at 37°C, the *E. coli*[pSL122] (containing the BamHI polynucleotide) and *E. coli*[pSL145] (containing the BamHI fragment with a deletion of orf3 to orf5) strains exhibit a strong yellow coloration, whereas *E. coli*[pBC SK^+] (containing the intact cloning vector) and *E. coli*[pSL127] (containing the BamHI fragment with a deletion of orf2 to orf5) are not colored.

This result demonstrates the involvement of the two genes orf1 (albA, SEQ ID No.1) and orf2 (albB, SEQ ID No.2) in the cyclodipeptide oxidase activity associated with the production of albonoursin.

EXAMPLE 6: Expression of the cyclodipeptide oxidase (CDO) enzyme activity in a heterologous host, *Streptomyces lividans*, transformed by introduction of orf1 and orf2 (albA and albB, SEQ ID No.1 and SEQ ID No.2).

The HPLC analysis of the supernatant from culturing the *Streptomyces lividans* TK21 strain transformed with the plasmid pSL142 containing the

BamHI polynucleotide from which orf3 (albC, SEQ ID No.3), orf4 (albD, SEQ ID No.4) and orf5 have been deleted, under the culture conditions described in example 4, demonstrate an absence of production of albonoursin and therefore indicates the involvement of orf3 and/or orf4 in the production of the diketopiperazine derivative.

However, the addition of cyclo(L-Phe-L-Leu) to the supernatant, under the standard conditions described in Gondry et al. (mentioned above), clearly results in the production of albonoursin. This suggests the involvement of orf3 and/or orf4 in the biosynthesis of the cyclodipeptide cyclo(L-Phe-L-Leu).

EXAMPLE 7: Demonstration of the production of cyclo(L-Phe-L-Leu) and cyclo(L-Phe-L-Phe) *in vivo* in a heterologous host, *Streptomyces lividans*, transformed by introduction of orf3 (albC, SEQ ID No.3)

In order to confirm the results described in example 6, orf3 (albC) and orf4 (albD) were cloned separately into the shuttle plasmid pUWL201, giving, respectively, pSL168 (orf3) and pSL159 (orf4), which were introduced into *Streptomyces lividans* TK21 according to the standard protocols described by Sambrook et al. (mentioned above) and Kieser et al. (mentioned above).

After culturing under the conditions described in example 4, the *S. lividans* [pSL168] and *S. lividans* [pSL159] culture supernatants were analyzed by HPLC, under the standard conditions described in example 4, in order to demonstrate the production of the cyclodipeptide cyclo(L-Phe-L-Leu). Since direct detection of this compound is difficult due to its low molar adsorption coefficient ($\varepsilon_{mol} \approx 100\ M^{-1}cm^{-1}$ at 254 nm) and to the complexity of the medium, the demonstration thereof was carried out after conversion of the cyclodipeptide to albonoursin (cyclo(ΔPhe-ΔLeu), $\varepsilon_{mol} = 25120\ M^{-1}.cm^{-1}$ at 318 nm), by addition of the CDO enzyme purified according to the method described in Gondry et al. (mentioned above).

The culture supernatants were filtered, and then incubated for 10 to 15 hours at 30°C with 4.1×10^{-3} enzyme units of purified CDO. Compared HPLC analysis of the culture supernatants, incubated or not incubated with CDO, was performed. The molecular mass of the metabolites produced was determined by mass spectrometry (Quattro II, Micromass).

The results are given in figure 6:

In the culture supernatant from *S. lividans*[pSL168], incubated in the presence of CDO, albonoursin (cyclo(ΔPhe-ΔLeu)) (peak at 40.5 min; λ_{max} = 318 nm; m = 256.4 Da) and cyclo(ΔPhe-ΔPhe) (peak at 44.1 min; λ_{max} = 338 nm; m = 290.3) (panel A) are detected;

No metabolite is detected:

- in the culture supernatant from *S. lividans*[psL168] in the absence of CDO (panel B);
- in the culture supernatant from *S. lividans*[pSL159] in the absence or in the presence of CDO (panel C);
- in the culture supernatant from *Streptomyces lividans* TK21 incubated in the presence of CDO: (panel D).

This result clearly demonstrates the involvement of orf3 (albC) in the production of the cyclodipeptides cyclo(L-Phe-L-Leu), the precursor for albonoursin, and cyclo(l-Phe-L-Phe), the precursor for a second metabolite, cyclo(ΔPhe-ΔPhe), produced together initially in *Streptomyces noursei* (Khokhlov A.S. et al., Tetrahedron Lett., 27, 1881 (1963)).

EXAMPLE 8: Demonstration of the production of cyclo-(L-Phe-L-Leu) and cyclo(L-Phe-L-Phe) *in vivo* in a heterologous host, *Streptomyces lividans*, transformed by introduction of the polynucleotide orf3-orf4 (albC-albD, SEQ ID No.3-SEQ ID No.4)

According to a variation of example 7, the polynucleotide orf3-orf4 (albC-albD) was cloned into the shuttle plasmid pUWL201, and the resulting plasmid, pSL167, was introduced into *Streptomyces lividans* TK21 according to the standard protocols described by Sambrook et al. (mentioned above) and Kieser et al.

(mentioned above).

HPLC analysis of the culture supernatant, treated in an identical manner to the method described in example 7, shows that, in the culture supernatant from *S. lividans*[pSL167] incubated in the presence of CDO, albonoursine (cyclo(ΔPhe-ΔLeu)) and cyclo(ΔPhe-ΔPhe) are detected and that no metabolite is detected in the culture supernatant from *S. lividans*[pSL167] in the absence of CDO, nor in the culture supernatant from *Streptomyces lividans* TK21 incubated in the presence of CDO.

These results demonstrate that the product of the orf5 gene does not participate directly in the biosynthesis of albonoursine, whereas ofr4, (albC), is necessary and sufficient to produce the precursor cyclodipeptides cyclo(L-Phe-L-Leu) and cyclo(L-Phe-L-Phe), in *Streptomyces lividans* as in *Streptomyces noursei*.

EXAMPLE 9: Overexpression of AlbA and AlbB from the polynucleotide albA-albB (SEQ ID No.1-SEQ ID No.2)

The vector pET-28a(+), which contains an N-terminal or C-terminal polyhistidine sequence (His tag), was used to construct an expression vector containing the polynucleotide albA-albB, in order to facilitate the purification of the recombinant protein. The shortest sequence of albA, encoding a polypeptide of 196 amino acid residues, was chosen. The gene amplification of the polynucleotide by PCR was carried out using oligonucleotide primers designed so as to include an NdeI (sense) and XhoI (antisense) cloning site.

The PCR conditions are as follows: initial denaturation at 94°C for 4 min, followed by 10 cycles of 1 minute at 94°C, 1 minute at 45°C and 1.5 minutes at 72°C, and then by 20 cycles of 1 minute at 94°C, 1 minute at 50°C and 1.5 minutes at 72°C, and a final polymerization reaction at 72°C for 10 min. The reaction products were digested with NdeI and XhoI, and the fragments were subcloned into the vector pET-28a, to give pSL150. The sequence of the insert was verified

by automatic sequencing (ABI PRISM Genetic analyzer, Perkin Elmer) using the DYEnamic ET terminator cycle kit (Amersham Pharmacia Biotech).

Standard expression conditions were used:

- growth temperature: 20°C,
- induction of expression: 0.6 mM IPTG
- induction time: 16 hours.

pSL150 was introduced into *E. coli* BL21 (DE3)-plysS. The strain was cultured in LB medium at 20°C until an absorbence of 0.6 was obtained, and then expression was induced. The culture was then centrifuged at 4190 g, at 4°C for 15 min. The cells were then resuspended in extraction buffer (100 mM Tris-HCl, pH 8.0, 1 µM phosphoramidon, 1 mM PMSF and 5% glycerol), and ground with an Eaton press. The protein extract was incubated in the presence of benzonase (25 U/ml) at 30°C for 10 min, and then centrifuged at 11 300 g for 15 min at 4°C. The enzyme activity was determined according to the standard assay described by Gondry et al. (mentioned above). One unit of enzyme activity is defined as the amount of enzyme that catalyzes the production of 1 µmol of product per minute, and the specific activity is expressed in units of enzyme per mg of proteins. The specific activity of the enzyme extract was increased by a factor of 50 (As = 2 U/mg) after purification of affinity chromatography (column: HiTrap chelating HP (1 ml), Amersham Pharmacia Biotech; equilibration of Ni^{2+} ions in 100 mM Tris-HCl buffer, pH 8.0, containing 0.5 M NaCl and 10 mM imidazole; elution: gradient of 0.3 to 1 M imidazole in 35 min; flow rate: 1 ml/min).

The sodium dodecyl sulfate/polyacrylamide gel (SDS-PAGE) (12%) analysis of the purified fraction of *E. coli*[pSL150] (figure 7) demonstrates the simultaneous presence of AlbA and AlbB in non-stoichiometric proportions, and demonstrates that AlbB is expressed as 2 isoforms (2 bands in SDS-PAGE). The analysis of purified AlbB by mass spectrometry and N-terminal sequencing of the polypeptide sequence indicates that

these two forms correspond to the products $AlbB_1$ and $AlbB_2$ from the 2 initiation codons identified in its sequence (cf. above).

EXAMPLE 10: *In vitro* conversion of cyclo(L-Phe-L-His) and cyclo(L-Phe-L-Leu) by recombinant AlbA-AlbB

The enzyme preparation purified according to the method described in example 9, incubated under the standard conditions as described by Gondry et al. (mentioned above), catalyzes the *in vitro* conversion of cyclopeptides to monodehydro and bisdehydro cyclodipeptides.

The purified enzyme preparation was incubated in the presence of the cyclo(L-Phe-L-His) substrates for 72 h at 30°C. The reaction products were analyzed by reverse-phase HPLC under the conditions described in example 4, and identified on the basis of their spectral characteristics and of their molecular mass confirmed by mass spectrometry. The reaction products are as follows:

- from cyclo(L-Phe-L-His): cyclo(ΔPhe-L-His) at (λ_{max} = 297 nm and m = 282 Da) and cyclo(ΔPhe-Δ-His) (λ_{max} = 338 nm and m = 280 Da),
- from cyclo(L-Phe-L-Leu): cyclo(ΔPhe-L-Leu) (λ_{max} = 297 nm and m = 258 Da) and cyclo(ΔPhe-Δ-Leu) (λ_{max} = 316 nm and m = 256 Da).

These results confirm that the enzyme preparation obtained from cloning the polynucleotide albA-albB into an expression vector introduced into a heterologous host catalyzes *in vitro*, the conversion of cyclodipeptides to α,β-dehydrogenated diketopiperazine derivatives.